

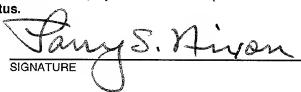
FORM PTO-1390 (REV 11-2000)	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER 36-1451
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5) <b>09/889512</b> Unknown
INTERNATIONAL APPLICATION NO. PCT/GB00/00440	INTERNATIONAL FILING DATE 18 February 2000	PRIORITY DATE CLAIMED 18 February 1999 18 June 1999
TITLE OF INVENTION		TRANSLATION
APPLICANT(S) FOR DO/EO/US APPLEBY Stephen Clifford.		

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

- ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
- ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
- ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
- ☒ The U.S. has been elected by the expiration of 19 months from the priority date (Article 31).
- ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2)).
  - ☒ is attached hereto (required only if not communicated by the International Bureau).
  - ☒ has been communicated by the International Bureau.
  - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
- ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
  - ☐ is attached hereto.
  - ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
- ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
  - ☐ are attached hereto (required only if not communicated by the International Bureau).
  - ☐ have been communicated by the International Bureau.
  - ☐ have not been made; however, the time limit for making such amendments has **NOT** expired.
  - ☐ have not been made and will not be made.
- ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
- ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
- ☐ An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 To 20 below concern document(s) or information included:

- ☐ An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98.
- ☒ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. 3.28 and 3.31 is included.
- ☒ A FIRST preliminary amendment.
- ☐ A SECOND or SUBSEQUENT preliminary amendment.
- ☐ A substitute specification.
- ☐ A change of power of attorney and/or address letter.
- ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821-1.825.
- ☐ A second copy of the published international application under 35 U.S.C. 154(d)(4).
- ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
- ☐ Other items or information.

U.S. APPLICATION NO. (If not previously filed) <b>09/1889512</b> US 09/1889512	INTERNATIONAL APPLICATION NO <b>PCT/GB00/00440</b>	ATTORNEY'S DOCKET NUMBER <b>36-1451</b>	
21. <input checked="" type="checkbox"/> The following fees are submitted:		<b>CALCULATIONS</b> PTO USE ONLY	
<b>BASIC NATIONAL FEE (37 C.F.R. 1.492(a)(1)-(5):</b> -- Neither international preliminary examination fee (37 C.F.R. 1.482) nor international search fee (37 C.F.R. 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO .....\$1000.00 -- International preliminary examination fee (37 C.F.R. 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO.....\$860.00 -- International preliminary examination fee (37 C.F.R. 1.482) not paid to USPTO but international search fee (37 C.F.R. 1.445(a)(2)) paid to USPTO.....\$710.00 -- International preliminary examination fee (37 C.F.R. 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4).....\$690.00 -- International preliminary examination fee (37 C.F.R. 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4).....\$100.00			
<b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>		\$      860.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 C.F.R. 1.492(e)).		\$      0.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE
Total Claims	12	-20 =	0
Independent Claims	5	-3 =	2
MULTIPLE DEPENDENT CLAIMS(S) (if applicable)			\$270.00
<b>TOTAL OF ABOVE CALCULATIONS =</b>			<b>\$ 1020.00</b>
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.			0.00
<b>SUBTOTAL =</b>			<b>\$ 1020.00</b>
Processing fee of \$130.00, for furnishing the English Translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 C.F.R. 1.492(f)).			0.00
<b>TOTAL NATIONAL FEE =</b>			<b>\$ 1020.00</b>
Fee for recording the enclosed assignment (37 C.F.R. 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. 3.28, 3.31). \$40.00 per property			40.00
Fee for Petition to Revive Unintentionally Abandoned Application (\$1240.00 - Small Entity = \$620.00)			0.00
<b>TOTAL FEES ENCLOSED =</b>			<b>\$ 1060.00</b>
		Amount to be: refunded	\$
		Charged	\$
a. <input checked="" type="checkbox"/> A check in the amount of \$1060.00 to cover the above fees is enclosed. b. <input type="checkbox"/> Please charge my Deposit Account No. 14-1140 in the amount of \$_____ to cover the above fees. A duplicate copy of this form is enclosed. c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 14-1140. A duplicate copy of this form is enclosed. d. <input checked="" type="checkbox"/> The entire content of the foreign application(s), referred to in this application is/are hereby incorporated by reference in this application.			
<b>NOTE: Where an appropriate time limit under 37 C.F.R. 1.494 or 1.495 has not been met, a petition to revive (37 C.F.R. 1.137(a) or (b)) must be filed and granted to restore the application to pending status.</b>			
<b>SEND ALL CORRESPONDENCE TO:</b>  NIXON & VANDERHYE P.C. 1100 North Glebe Road, 8 <sup>th</sup> Floor Arlington, Virginia 22201-4714 Telephone: (703) 816-4000			
		 SIGNATURE	
		<b>Larry S. Nixon</b> NAME	
<b>25,640</b> REGISTRATION NUMBER		<b>July 18, 2001</b> Date	

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of

**APPLEBY**

Atty. Ref.: **36-1451**

Serial No. **Unknown**

Group:

National Phase of: **PCT/GB00/00440**

International Filing Date: **11 February 2000**

Filed: **July 18, 2001**

Examiner:

For: **TRANSLATION**

\* \* \* \* \*

**July 18, 2001**

Assistant Commissioner for Patents  
Washington, DC 20231

Sir:

**PRELIMINARY AMENDMENT**

Prior to calculation of the filing fee and in order to place the above identified application in better condition for examination, please amend the claims as follows:

**IN THE CLAIMS**

Please substitute the following amended claims for corresponding claims previously presented. A copy of the amended claims showing current revisions is attached.

3. (Amended) Apparatus according to claim 1, wherein multiple said events are referenced to a common said alternation class.

4. (Amended) Apparatus according to claim 1, wherein multiple said alternation classes are referenced to a common said set of language-independent roles.

5. (Amended) Apparatus according to claim 1, wherein the alternations of an alternation class constitute a set of language-independent rules, and the respective sets of language-dependent alternation classes for all the languages of the apparatus together constitute said means for storing sets of language-independent roles.

**APPLEBY**

Serial No. **Unknown**

9. (Amended) A method according to claim 7, wherein, in said step of storing a plurality of words, multiple said words representing a said event are referenced to a common said alternation class.

**REMARKS**

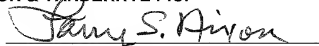
Attached hereto is a marked-up version of the changes made to the claims by the current amendment. The attached page is captioned "**Version with markings to show changes made.**"

The above amendments are made to place the claims in a more traditional format.

Respectfully submitted,

**NIXON & VANDERHYE P.C.**

By:



**Larry S. Nixon**

Reg. No. 25,640

**LSN:Imy**

1100 North Glebe Road, 8th Floor  
Arlington, VA 22201-4714  
Telephone: (703) 816-4000  
Facsimile: (703) 816-4100

**VERSION WITH MARKINGS TO SHOW CHANGES MADE**

3. (Amended) Apparatus according to [either preceding claim] claim 1, wherein multiple said events are referenced to a common said alternation class.

4. (Amended) Apparatus according to [any preceding claim] claim 1, wherein multiple said alternation classes are referenced to a common said set of language-independent roles.

5. (Amended) Apparatus according to [any preceding claim] claim 1, wherein the alternations of an alternation class constitute a set of language-independent rules, and the respective sets of language-dependent alternation classes for all the languages of the apparatus together constitute said means for storing sets of language-independent roles.

9. (Amended) A method according to [any either claim 7 or claim 8] claim 7, wherein, in said step of storing a plurality of words, multiple said words representing a said event are referenced to a common said alternation class.

1  
TRANSLATION

This invention relates to automatic language translation.

Machine language translators accept input text in a first natural language (the source language) and generate corresponding output text in a second natural language (the target language). Such translators may be classified into two types; those which use a set of translation rules for each possible pair of source and target languages, and those (relatively rare) interlingual systems which translate from the source language into a language independent (interlingual) form, and then from this language independent form to the target language.

In the system described in our earlier application number PCT/GB98/02389, rules specifying the complements which each verb of all source and target languages could take were present, and were stored with pointers from the corresponding verb entries in a lexical database. These rules also specified the mapping between the complements and the roles (e.g. agent or patient) corresponding to them.

The roles were assigned in a relatively simple way, with the subject of the verb always the active role (agent) and the object the passive role (patient). Abstraction rules then dealt with the necessary changes to the role in unusual cases. Complements attached by prepositional phrases would not have roles; these were assigned by abstraction rules.

The rules needed to be hand written, and since this was required on the order of one per verb per language, the effort was considerable and the results were not consistent.

According to a first aspect of the present invention, there is provided an apparatus for translating text from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the apparatus comprising:

means for storing sets of language-independent roles;

means for storing, for each said language,

a respective set of language-dependent alternation classes, each alternation class being referenced to a said set of language-independent roles and comprising a set of alternations, and

a plurality of words, each word representing a said event and being referenced to an alternation class of its respective set of language-specific alternation classes,

each alternation of an alternation class comprising, in a respective order, the roles of said set of language-independent roles to which that class is referenced and listing the correspondence between those roles and complements by which each event referenced to that alternation class can validly be represented in that language;

means for locating, in said text, a phrase comprising one or more words representing an event, and the complements representing roles associated with that event;

means for representing said phrase in a language-dependent semantic structure;

and

means for replacing said language-dependent semantic structure with an indication of the language-independent roles represented by said complements using said alternations.

Preferably, when said events are defined by a verb in said source language, said means for replacing are arranged to locate a preposition occurring with said verb, and to replace said preposition with an appropriate role.

Preferably, multiple said events are referenced to a common said alternation class.

Preferably, multiple said alternation classes are referenced to a common said set of language-independent roles.

Preferably, the alternations of an alternation class constitute a set of language-independent rules, and the respective sets of language-dependent alternation classes for all the languages of the apparatus together constitute said means for storing sets of language-independent roles.

According to a second aspect of the present invention, there is provided a method of setting up a machine translation system comprising:

a first stage of establishing, for each language, a respective plurality of language-dependent alternation classes, each alternation class comprising a plurality of alternations derived using a common set of language-independent role data shared by words of each language with a common meaning, and a respective set of word entries.

each word entry being referenced to a said respective language-dependent alternation class; and

a second stage of creating, for each said alternation class and for each word entry referenced to that alternation class, a corresponding set of word entries identical to the word entry referenced to that alternation class, each word entry of said set being referenced to a different one of the alternations of that alternation class.

According to a third aspect of the present invention, there is provided a method of translating text from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the method comprising the steps of:

storing sets of language-independent roles;

storing, for each said language,

a respective set of language-dependent alternation classes, each alternation class being referenced to a said set of language-independent roles and comprising a set of alternations, and

a plurality of words, each word representing a said event and being referenced to an alternation class of its respective set of language-specific alternation classes,

each alternation of an alternation class comprising, in a respective order, the roles of the set of language-independent roles to which that class is referenced and listing the correspondence between those roles and complements by which each event referenced to that alternation class can validly be represented in that language;

locating, in said text, a phrase comprising one or more words representing an event, and the complements representing roles associated with that event;

representing said phrase in a language-dependent semantic structure; and

replacing said language-dependent semantic structure with an indication of the language-independent roles represented by said complements using said alternations.

Preferably, said events are defined by a verb in said source language, and said replacing step is arranged to locate a preposition occurring with said verb, and to replace said preposition with an appropriate role.

Preferably, in said step of storing a plurality of words, multiple said words representing a said event are referenced to a common said alternation class.



Preferably, in said step of storing a respective set of language-dependent alternation classes, multiple said alternation classes are referenced to a common said set of language-independent roles.

Since the presence of prepositions and other verb form irregularities is captured in a relatively small number of alternation records, a relatively small number of parsing and abstraction rules to replace such prepositions and other features on detection of their occurrence can be employed.

Other aspects and preferred embodiments are as described in the following description and claims.

Embodiments of the invention will now be illustrated, by way of example only, with reference to the accompanying drawings, in which:

Figure 1 is a block diagram of the language translation apparatus according to a first embodiment;

Figure 2 is a block diagram showing in greater detail the processes present in a client terminal forming part of the embodiment of Figure 1;

Figure 3 is a block diagram showing in greater detail the processes present in a server forming part of the embodiment of Figure 1;

Figure 4 is a block diagram showing in greater detail the subprocesses present within a translation process forming part of the embodiment of Figure 3;

Figure 5 is an illustrative diagram showing the formats through which text passes during the translation process of the embodiment of Figure 1;

Figure 6 is a block diagram showing the databases maintained within the server of Figure 1;

Figure 7 is a schematic diagram illustrating the word structure produced after text pre-processing in the embodiment of Figure 1;

Figure 8 is a diagram illustrating the entity/relationship semantic structure produced after parsing in the embodiment of Figure 1;

Figure 9 is a flow diagram showing schematically the operation of the server of the embodiment of Figure 1;

Figure 10 is a diagram illustrating the data stores present in the embodiment;

Figure 11 is a diagram illustrating the relationship between records in the embodiment;

Figure 12 is a flow diagram illustrating a compilation phase of writing records to each word store of Figure 10;

Figure 13 is a diagram illustrating the relationship between records during the process of Figure 12;

Figure 14a illustrates a pre-abstracted, language-specific structure, and

Figure 14b illustrates the structure after application of an abstraction rule,

Figure 15 is a flow diagram illustrating the abstraction rule operation;

Figure 16 (comprising Figures 16a and 16b) is a flow diagram illustrating the process of generating the data used in the present embodiment; and

Figure 17 illustrates a role set record of the embodiment.

### Background to Embodiment

For ease of reading, features of PCT/GB98/02389 are reiterated here; the whole of the description is incorporated herein by reference.

Referring to Figure 1, the present invention may be employed by a client terminal 100 connected via a telecommunications network 300 such as the Public Switched Telephone Network (PSTN) to a server computer 200. The terms "client" and "server" in this embodiment are illustrative but not limiting to any particular architecture or functionality.

The client terminal comprises a keyboard 102, a VDU 104, a modem 106, and a computer 108 comprising a processor, mass storage such as a hard disk drive, and working storage, such as RAM. For example, a SUN™ work station or a Pentium™ personal computer may be employed as the client terminal 100.

Stored within the client terminal (e.g. on the hard disk drive thereof) is an operating control program 110 comprising an operating system 112 (such as Windows™), a browser 114 (such as Windows Explorer™ Version 3) and an application designed to operate with the browser 114, termed an applet, 116. The function of the operating system is conventional and will not be described further. The function of the browser 114 is to interact, in known fashion, with hypertext information received from the server 200 via the PSTN 300 and modem 106. The browser 114 thereby downloads the applet 116 at the beginning of the communications session, as part of a hypertext document from the server 200. The function of the applet 116 is to control the display of received information, and to allow the input of information for uploading to the server 200 by the user, through the browser 114.

Referring to Figure 3, the server 200 comprises an operating program 210 comprising an operating system 212 such as Unix™, a server program 214 and a translator program 216. The operating system is conventional and will not be described further. The function of the server program 214 is to receive requests for hypertext documents from the client terminal 100 and to supply hypertext documents in reply. Specifically, the server program 214 initially downloads a document containing the applet 116 for the client terminal 100. The server program 214 is also arranged to supply data to and receive data from the translator program 216, via, for example, a cgi.bin mechanism.

The function of the translator program 216 is to receive text from the client terminal 100 via the telecommunications network 300 and server program 214; to interact with the user as necessary in order to clarify the text; and to produce a translation of the text for supply back to the user (in this embodiment).

Figure 4 shows the component programs of the translator 216. It comprises a number of sections; one for each language, of which only a first section 220, relating to a first language (LANG1) and a second section 230 relating to a second language (LANG2), are shown for clarity. Each language section comprises the following subprograms or modules:

- 1) A text pre-processor (221, 231)
- 2) A source language parser (222, 232)
- 3) A source language abstractor (223, 233)
- 4) A target language de-abstractor (224, 234)
- 5) A target language generator (225, 235)
- 6) A target language text post-processor (226, 236)

The functions of each of these components will be discussed in greater detail below.

Figure 5 illustrates the stages of translation according to this embodiment.

A source language text document (stage A) is received by the translator from the client terminal 100.

After operation of the text pre-processor stage (221), the result is an expanded source language text document (stage B). The operation of the pre-processor is to replace contracted forms of words (such as "he's" in English, or "j'ai" in French) with their non-contracted forms.

After operation of the source language parser 222, the result at stage C of Figure 5 is a language-specific semantic structure which represents the input text as an encoded entity-relationship graph, where the entities are semantic categories corresponding to the words (in other words, identifying the nouns, verbs and so on), and the relationships are data relating the entities together (e.g. to indicate those which are the subjects or objects of others).

After operation of the source language abstractor 223, the result at stage D is a further semantic structure D, similar to the language specific semantic structure produced at stage C but indicating additional relationships and data which substitute the language-specific meanings of some of the structures represented within the semantic structure C with abstracted structures.

For example, a phrase such as "My name is David" input as source language text could be represented within a parsed semantic structure by data indicating ownership of the name by the individual first person, and an attribute of the name being that it is "David". This is a grammatically correct expression, from which French or German text could be generated by a suitable generator such as 235.

However, whilst grammatical French or German would be produced, the meaning would be unclear, since in French the equivalent phrase is "I call myself" ("je m'appelle") and in German the equivalent phrase is "Ich heiÙe" (which is equivalent to "I am called" in English, but for which English lacks a corresponding verb). Accordingly, the source language abstractor 223 recognises, within the parsed semantic structure, the occurrence of structures which are not directly translatable, such as structures involving personal names in this example, and replaces those structures with additional data representing them.

Accordingly, the abstracted semantic structure produced at stage D of Figure 5 corresponds to a representation of the input text but with the replacement of specific constructs which are known not to meaningfully translate into one or more other languages (whether or not those languages are represented by sections within the translator 216).

The abstracted semantic structure produced at stage D is an interlingual form which is unambiguous in relation to each of the target languages which the system is capable of translating into. That is to say that the interlingual form corresponds uniquely with a language-specific semantic structure in each of the target languages.

The abstracted semantic structure, or one of the abstracted semantic structures, produced by the abstractor in stage D is then passed to the de-abstractor 234 of the target language, which comprises a series of rules which test for the presence of the additional structures inserted by the language abstractor 223, and translate them into the form used in the target language. For instance, in the example given above, the abstracted naming operation would be converted, in French, into "je me appelle" (I call myself). The result is then, at stage E, a semantic structure equivalent to the language-specific semantic structure at stage C but in which the semantic substructures corresponding to phrases or expressions in the input text which would give rise to translation difficulties have been replaced by appropriate substructures in the target language. This structure forms the input to the target language generator 235, which generates a corresponding target language output text (stage F), and therefore applies the reverse process to the parsers 222, 232.

Finally, the generated output text at stage F is contracted by the text post-processor 236 which takes the generated text and contracts relevant parts of it. In the above example, "je me appelle David" would be contracted to "je m'appelle David". Other minor text processing operations, such as adding capital letters at appropriate places (for example at the beginning of each sentence), and providing the correct spacing between words, are also carried out.

Referring to Figure 6, the server 200 stores data for use by the parser and abstractor in each language. This data comprises, for each language, a grammar rules database (227, 237) and an abstraction rules database (228, 238). Also present is a multilingual lexical database 240. The lexical database 240 stores an entry for each concept represented by a word in any language represented within the translator program, the entry pointing to corresponding word entries, in word stores (see Figure 10), in each of the languages within which a word exists which expresses that concept. The word entries give for each word of the text in the language concerned; the type of lexical element represented by the word (e.g. whether it is a noun, a verb, a pronoun, an adjective and so on); data on the manner in which the word is inflected, if at all, in each language, and various other data.

The grammar rules stored within each grammar rules database (227, 237) represent, for the corresponding language, the ways in which words of that language may be combined.

The operation of this embodiment will now be disclosed in greater detail with reference to Figures 7-11.

Referring to Figure 9, in a step 402, text is received from the client terminal 100. In a step 404, the input text is expanded. As a first step, the start and end of each possible word in the text is located by detecting spaces and punctuation, so as to result in a stream of possible words. As a second step, any contracted words (such as "j'ai" in French") are expanded to replace them with full words (in that example, "je ai"). At the same time, the text pre-processor locates and flags special text items such as proper names, dates, times, sums of money and so on.

At this stage, there may be several possible expanded strings of words that could match each contracted string of word. All such possibilities are retained as alternatives.

Next, each word is looked up in the lexical database 240. At this stage, words which are not recognised but are closely matched to others in the source language (that is, the language of the input text) are replaced by all those for which they are a close match, as in the manner of a conventional spell checker.

If, after spell checking, any words have not been recognised (step 405) then a query is transmitted back to the user, comprising a text message saying, for example, "The word (unrecognised word) has not been recognised. Please check the spelling, and resubmit this word or a synonym". This query is then transmitted to the client terminal 100 in step 406.

The result of this pre-processing is therefore that the expanded text (stage B of Figure 5) is no longer necessarily a linear sequence of words but may, as shown in Figure 7, comprise a network or lattice of words.

Figure 7 indicates such a network in which the second word, originally B, has been replaced by two possible alternatives (either alternative spellings or alternative expansions) B1 and B2, and the third word C has been replaced by three possible alternatives C1, C2 and C3. There are thus now six possible routes through the network of words.

The text of each word in the network is now replaced by a reference to the corresponding entry in the lexical database 240. If a single word (such as "bank" in English) has two different entries in the lexical database 240 corresponding to different meanings (which would be translated into different words in a target language), the word is replaced by each possible entry in the lexical database 240. For convenience,

rather than using references to the entries in the lexical database, the syntactic category information for each word (i.e. whether it is a noun, verb etc.) may be retained within the network, and a table relating each network position to the corresponding entry in the lexical database 240 may be separately stored for later use.

On each occasion where a single word in the source language is given as the translation of several different lexical entities in the database 240 (corresponding to several different words in one or more of the target languages), a reference to each of these is included within the processed text lattice of Figure 7.

Further details of parsing are given below.

Next, the network of nodes (each corresponding, as noted above, to one of the entries in the lexical database 240 and being represented by the syntactic category of that entry) is processed by the source language parser program, which, for each word, applies the rules within the grammar rules database 227 which are applicable to words of that type.

Thus, for example, referring to Figure 8, suppose that the English text contained the phrase "the dog saw the cat". The word "the" is the definite article, and a rule within the grammar rules database 227 indicates that it can be followed by the noun to which it refers. Thus, the circle D1 indicating the first occurrence of determiner "the" is linked by this rule to the next circle N1, representing the following noun "dog", and the circle D2, representing the second occurrence of determiner "the" is linked by this rule to the circle N2 for the following word, which is the noun "cat".

The rule for the active form of the verb "to see" indicates that the verb may be preceded by the seeing "agent" entity (in this case "the dog") and followed by the patient entity (in this case "the cat").

Thus, after parsing, the parsed semantic structure (stage C of Figure 5) is represented, for each sentence of the input text, by one or more structures comprising references to entries in the lexical database 240 (the circles in Figure 8) and pointers linking them together (the lines in Figure 8). In the PROLOG computer language, the topological structure of Figure 8 may be represented as

[  
 $A \wedge \text{det}(\text{def}, s, \_ , \text{third}), A \wedge \text{e}(\text{dog}, []) , P \wedge \text{det}(\text{def}, s, \_ , \text{third}), P \wedge \text{e}(\text{cat}, []) ,$   
 $E \wedge \text{event}(\text{see}, \text{fin}, \text{past}, []) , E \wedge A \wedge r(\text{agent}, []) , E \wedge P \wedge r(\text{patient}, [])$   
 ]

In the foregoing, it will be noted that the unifying variables A and P are the links which unify the first occurrence of "the" with "dog" and the second occurrence of "the" with "cat". The verb "see" is linked by an agent relationship and a patient relationship with the terms linked by the relationship A (i.e. "the dog") and the terms linked by the relationship P (i.e. "the cat").

The verb is recorded as an event ("event"), and is linked to the lexical entry in the lexical database 240 for the word "see" and is indicated to be the finite form ("fin") in the past tense ("past").

The word "the" is recorded as a determiner, being the definite article ("def"), single rather than plural form ("s"), having neutral gender ("\_") and referring to the third person ("third"). The terms for "dog" and "cat" are indicated to be entities ("e"), and have a reference to the corresponding word entry in the lexical database 240.

Thus far, other than the target-language dependency, the parser is not dissimilar to known, technically and commercially available products. Further information on suitable chart-parsing techniques which may be used will be found in James Allen, "Natural Language Understanding", 2nd Edition, Benjamin Cummings Publications Inc., 1995.

In other respects, the parser may be as described in PCT/GB98/02389.

Having thus parsed the text (step 410), the abstractor 223 then accesses the abstracting rules database 228 to locate those source language phrases which may give rise to translation difficulties. The abstraction process is recursive, insofar as once one abstraction rule has been applied to the parsed text, the entire set of abstraction rules is referred to again when processing the partially abstracted text to identify another abstraction rule to be applied, repetitively until none of the abstraction rules in the set can be applied.

Thus, in step 412, the abstractor 223 tests each structure generated by the parser, and where one or more of the abstraction rules is applicable, converts the detected structure to the alternative form recorded within the rule. As explained, this test is recursive such that the same rule may be applied at different stages of an



abstraction process in which a structure generated by the parser is converted to the interlingual structure.

After operation of the abstractor 223, the ideal result should be a single, complete interlingual structure. If the structure is incomplete (that is to say, it was not possible to relate together all the words using the grammar and the abstraction rules) then successful translation will not be possible. If more than one possible structure is produced, then the input text is considered ambiguous since it could result in more than one possible translation in at least one of the target languages. If either of these conditions is met (step 414), a query is transmitted to the user (step 406).

In greater detail, the problematic points within the semantic structure, corresponding to incomplete or ambiguous meanings, are located, and the portions of the input text relating to these are formulated into a message and transmitted back to the user for display and response by the applet 116, with a query text which may for example say "the following text has not been understood/is ambiguous."

In a preferred version of the present embodiment, the de-abtractor and generator 224, 225 corresponding to the input (source) language are employed (as described in greater detail below) to generate a source language text for each possible semantic structure where two or more such structures exist, and the query also includes these texts, prefixed with a statement "one of the following meanings may be intended, please indicate which is applicable:"

In this case, the message transmitted to the user in step 406 comprises a form, with control areas which may be selected by the user at the client terminal 100 to indicate an intended meaning for the ambiguous words or phrases detected within the input text.

If no such ambiguities are detected, or after all such ambiguities are resolved (step 414), the single, unified, interlingual semantic structure produced by the abstractor 223 is then passed to the target language de-abtractor 234 for the or each target language into which the text is to be translated. The de-abtractor 234 accesses the abstracting rules database 238 and, on detection of any of the substituted forms (for example "I sit") substitutes the normal form for the target language (in this case, "I sit myself" in French or "I am sitting" in English). The de-abstracted structure is then more idiomatically correct in the target language than was the semantic structure produced by the parser.

Next, in step 418, the target language generator program 235 accesses the target language grammar rule database 237 and the lexical database 240 and operates upon the de-abstracted semantic structure to generate output target language text.

The operation of the generator is essentially the reverse of that of the parser; briefly stated, it operates a chart-parsing algorithm (of a type known of itself) to take the components of the target language semantic structure generated by the de-abstractor, look up the applicable rules in the target language rules database 237, and assemble the corresponding words located from the lexical database 240 into a string of text ordered in accordance with the grammar rules, until a single stream of text which utilises all components of the semantic structure and obeys the grammatical rules is located.

After generating the output text stream, the text is post processed (step 420) to add a space before each word; capitalise the first letter in a sentence; add a full stop after the last word; contract any phrases (such as "je ai") which are capable of contraction; and reproduce any special forms of text (such as dates, amounts of money, and personal names), as appropriate for the target language.

The resulting formatted text is then formulated into an HTML (or text, or other suitable format) page, which is transmitted back to the user at the client terminal 100 in step 422.

On receipt of the translation result at the client terminal 100, the page is displayed via the browser 114 and may be converted and stored for subsequent word-processing by the user.

#### First Embodiment

In English, and in many other languages, a phrase involving a verb may have several different possible word orders, each of which is referred to here as an "alternation". This embodiment provides improved rules for dealing with the alternations which are associated with words (particularly verbs) which can take complements in several different orders (alternations). A description of alternations in English is to be found in "English verb classes and alternations", Beth Levin, Chicago Press 1993, ISBN 0226475336.

In English, many verbs have a prepositional phrase as a complement; for example; the verb "give" may have a noun phrase and a prepositional phrase as compliments, as in the example "I give [the book] [to the girl]". The preposition may not be required in the equivalent phrase in other languages. For example, in the English

phrase "they look for the ball", the preposition "for" is not represented in the French equivalent "ils cherchent la balle".

Others have a preposition-like particle attached - for example "bring in", in which the word "in" has no meaning except to modify the meaning of "bring".

Many verbs of attitude (i.e. verbs expressing states of mind) may have a reversible form. For example, the statement in English "I like the book" is equivalent in meaning to "the book pleases me", although there will in some cases be a subtle shift of emphasis. Both would be translated in Spanish, for example, as "el libro me gusta".

In each case it will be seen that the rules governing the use of the verb and the surrounding word order are quite language-specific, and therefore will need to have associated abstraction rules and de-abstraction rules. Unfortunately, to write separate abstraction rules for each verb is an enormous task for each language separately, and leaves open the risk that rules may be missed. It also leads to ad hoc and unsystematic development of the abstraction rules.

This embodiment therefore provides a new method of handling alternations (particularly for verbs), and new methods of use thereof.

#### Data Structures

Referring to Figure 10, the data structures employed in the present embodiment will now be described.

Referring to Figure 10, the lexical database 240 comprises, for each language, a list 1241, 1242 of word entries or records. Each word record in each language list points to a concept or meaning record entry in a language independent meaning store 1243. Each record in the store 1243 contains meaning (i.e. semantic information) relating to a concept, described by the word entries which point to that record.

One suitable structure for such a meaning store (lexicon) is given in the WordNet (TM) lexical database, available from Princeton University, Princeton, New Jersey, USA or MIT Press Five Cambridge Center, Cambridge, MA USA, details of which are at <http://www.cogsci.princeton.edu/~wn/>.

Thus, word records in the word list stores 1241, 1242 of different languages are indirectly linked, in that they point to a common entry in the semantic lexicon 240, which is related to the meaning expressed by the words.

Referring to Figures 10 and 11, each word store (1241, 1242) in a given language of the lexical database 240 contains a respective plurality of word entries (LEX1, LEX2, LEX3, LEX4), each word entry contains a pointer to an alternation class

record (702, 704, 706) in a corresponding one of a plurality of language-specific alternation class stores 1238, 1228 provided within respective grammar rules stores 228, 238 corresponding to each of the languages to be used.

The relationship is a many-to-one mapping. That is to say, many verbs (of the order of several thousand in English) map onto a relatively small number of alternation classes (of the order of 200 in this embodiment in English). Each word entry in the word store 1241, 1242 maps onto to only one alternation class record per language. Several lexical entries will share the same alternation class record. The significance of the alternation class records will be explained below.

In each of the alternation record stores 1228, 1238 (each said store being associated with a respective language), each alternation class record 702-706 is linked to one or more alternation records 708-722 by pointers.

#### Compiling the Word Entries

In this embodiment, the contents of the word stores 1241, 1242 are not fully populated until the apparatus is to be used. This results in a saving of memory space, since those word stores for languages which are not to be used in translation require less memory.

Accordingly, referring to Figures 12 and 13, a first set of uninflected word entries (WORD 1 of Figure 13 for example) remain resident in the store 1241 at all times. Where the word is one which can take multiple alternations, the word entry is linked by a pointer to an alternation class record (ALT CLASS of Figure 13) in the alternation store 1238 for the language concerned. This record is linked to the alternation records (ALT 1, ALT 2 of Figure 13) of the alternations which that word (and others sharing its alternation class) can take.

A program, operated prior to translation, performs the process of Figure 12. In a step 1402, the class record for a word is read, and a first alternation record is selected in a step 1404. In a step 1406, the program creates a new lexical entry (WORD 2) for the same word as that for WORD 1, which incorporates a pointer to the alternation class record, and stores the alternation as a list of possible complements in an order.

If there are more alternation records (step 1408), the process of step 1404 on is repeated for the next alternation (step 1410). If not, the next pre-existing word record in the store 1241 is selected (step 1414) and the process of step 1402 is repeated until all word entries have thus been expanded to an entry for each alternation (step 1412).

Referring to Figures 10 and 17, also provided in this embodiment is a role set store 740 storing a plurality of role set records, one of which is indicated (as 730) in Figure 17. Each role set record stores a plurality of role data, shown as R1, R2 and R3 in Figure 17. In the present embodiment, there are on the order of 15-20 role set records in total in the role set record store 740.

The entry for each event concept in the meaning store 1243 (e.g. each entity corresponding to an event or verb in a word store 1241, 1242) or other concept which can take an alternation in some language, is linked by a pointer to one of the role set records 730. It will therefore be clear that, since there are many thousands of such verb or event records in each word store of the lexical database 240, there is many-to-one mapping of lexical database entries to role set records.

Each of the alternation class records in the alternation class store 1241, 1242 for a language is linked by a pointer to one of the role set records 730. Since the number of role set records is substantially smaller than that of alternation class records in each language (e.g. by an order of magnitude), this too is a many-to-one mapping.

The significance of the data stored in each record will now be explained.

For each word representing an event in a language (e.g. a verb in English) the lexical database 240 stores a record for one or more semantic concepts which correspond to that word. For example, the word "give" in English may be associated with several different concepts, including "give to" and "give up". Each of these provides a different meaning for the word.

Within the word list for each language, separate entries are provided for each verb meaning, and in particular for each meaningful verb/preposition pairing. Thus, the verb "look" in English has one entry, but then "look for", "look at" and so on have separate entries, pointing to different concept records in the lexical database 240.

Associated with each such event concept, then, are one or more roles; that is to say, people or objects taking part in the event. For example, corresponding to the phrase "they looked for the ball", are an active role or agent ("they" - the people who looked) and a passive role or patient ("the ball" the thing that was looked at). Other events may involve more parties, for example, "[she] gave [it] [to him]" involves a donor, a recipient and an object. These roles are the same regardless of the word order employed (e.g. "he was given it by her") or the verb form used (e.g. "the book pleases me" and "I like the book"). They may vary slightly between languages, since in some languages certain roles may be inferred, but will be similar across all languages. Thus,

the role set record stores for a concept all roles which might be used in any language for that concept. They do not necessarily correspond to the subject and object of a verb, since these vary between active and passive forms of a verb.

Many different events (i.e. verbs) can be described using the same roles. Thus, according to the present embodiment it has been determined that a relatively small number of role sets (each represented by a role set record 730) can be provided, one or other of which will provide the necessary set of roles for all events in any language (see e.g. M A K Halliday "An introduction to Functional Grammar" (2nd Ed, 1994, Edward Arnold), ISBN 0340574917). The number of roles in each role set record, and their identities, differ from record to record. The role set records therefore represent a language-independent data structure linked to those concept records of the lexical database records which are for event concepts. Examples of (almost all the commonest) roles are

- For material "action" - agent and patient. This is the largest category with verbs of most actions.
- For "behavioral" events - a behavior. These are generally physiological events such as yawning.
- For "perception" and thinking events - sensor and phenomenon. Examples are Seeing, hearing, feeling.
- For "verbal" events - sender, message, and recipient. Examples are Saying, writing etc.
- For "relational" (stative) verbs - carrier, attribute or token, and value, depending on the kind of verb. For example, "the house is big", "John is the leader"
- For "existential" (there is/there are) - existent (i.e. that which exists).

Associated with each role in the role set in this embodiment is a restriction (not shown), the use of which will be discussed further below.

Each alternation class record 702-706 contains a record of the language it is relevant to, together with a name field naming the class record (for example, "reflexive verb"), and is linked by pointers to its role set record and alternation records.

Each alternation record 708-722 comprises a name field; a pointer to the alternation class record to which it is linked; the syntactic category (e.g. verb) of the alternation; and a list of terms each of which maps a role (which is one of those listed in the role set record associated with the alternation class to which the alternation record belongs) to a syntactic category (e.g. noun, preposition, and so on). Conveniently, the order of the mapping fields is significant. For a verb, the first mapping field is taken to indicate the subject of the verb (when it is present in the active form) and the remaining mapping fields indicate the complements of the verb in order.

Of the names for each alternation, one alternation of each alternation class is always named "normal"; this alternation will specify the word order most commonly used in the language concerned. The names of the other alternations in each alternation class will specify the conditions under which that alternation is used; for example, "polite", "formal", "stressed".

The declaration of an alternation (in PROLOG) may for example be as follows:

```

alternation(hold_onto_relate, relate, normal, eng,
  v:[],
  [
    agent-> np:[case = nom],
    patient-> pp:[pform = onto]
  ]
).

```

Here 'hold\_onto' is the name of the alternation class, 'relate' is the name of the role set record, 'normal' is the name of this alternation within the alternation class, 'eng' is the language code 'v:[]' is the syntactic category (verb).

Where an alternation specifies a prepositional phrase, the identity of the preposition is stored in the alternation (above, for example, "onto").

In some cases, where a particle such as the preposition-like particle "in" in the phrase "bring it in" can accompany the word, it is given a special "null" role, indicating it has no separate semantic significance. On compilation, this role is listed in the word record created from the alternation, and is hence required to be present if that alternation is to be detected during parsing, but no semantic terms are thereby introduced during translation.

Thus, the data stored in a set of alternations (for example, the set of different alternations in which the verb "hold" can be used with a preposition "onto" to express the concept "hold onto" in English) can be used to locate, within a given phrase involving that verb, which of the surrounding words or phrases occupies which role in relation to the verb.

This information can then be used to re-generate a corresponding phrase for the same concept expressed in the target language, since the set of roles is defined (in a language independent fashion) by the role set record to which both the source and target language alternation classes point, and from which both were derived.

In use, the present embodiment operates as follows.

#### Parsing

In this embodiment, parsing is performed as described above, to generate the language-specific parsed semantic structure.

During parsing, as mentioned above, each word in the expanded source language text document is looked up in the word store 1241 of the source language.



The word is then replaced by a reference to the word entry or entries corresponding to it, for subsequent use.

Where the word is a verb (for example, "give") which can be used in several different senses, several different entries will be found in the word store, (e.g. corresponding to "give to", "give up" and so on). Each of these different entries expresses a different concept, and therefore points to a different concept entry in the lexical database 240. Each also points to an alternation class record; however, two or more of these entries may point to a common alternation class record.

Further, where the word can take several alternations, a separate word entry for each alternation will have been created, as described above.

Thus, after all words have been looked up in the source language word store 1241, the parser uses the orders of the possible complements defined in the word records, together with the rules stored in the rule store, to attempt to create paths through the word lattice of Figure 7. Since only one of the alternations will actually be present, those word entries corresponding to alternations which have a complement order other than that detected to be present will be rejected during parsing.

Thus, after operation of the parser, the parsed semantic structure will include one or more identified alternations for each event term located in the input document, the alternations being identified where the syntactic categories surrounding the identified event in the source document match those in the order specified in the alternation.

The order of the role to syntactic category maps in the alternation record give a canonical ordering for the complements. The grammar has the power to vary this order for say, passives and relative clauses.

### Abstraction

At this stage, referring to Figure 14a, the semantic structure will still include any prepositions originally present; for example, the phrase "to the girl" will be identified as the patient entity in the phrase "he gave the ball to the girl", with "to" identified as a preposition. Also, in the case of verbs with prepositions and some other types of verb (for example, verbs in the passive form) the roles identified during parsing will not be language independent.

Accordingly, an abstracting rule is provided which, in the abstracting phrase, identifies each word term in the parsed semantic structure (step 1002), looks up the corresponding word record, and from that, accesses the corresponding alternation class

record (step 1004), and thence the alternation record (step 1006) corresponding to the alternation used to generate that word record.

If the alternation indicates that a preposition phrase is present in the source language (step 1008), then the abstraction rule deletes the entry for the preposition from the parsed semantic structure (step 1010), so that instead of pointing to the prepositional phrase "to the girl" as the object (or some other language dependent role), the event term points to the phrase "the girl" which followed the preposition.

Finally, in step 1012, data recording the language-independent role assigned to that prepositional phrase in the alternation record (here, "patient", as shown in Figure 14b) is assigned to the phrase which followed the preposition. The abstraction rule then proceeds in similar fashion until all terms in the parsed semantic structure are processed.

Other word forms where the assignment of complements to roles is language dependent can likewise be detected and amended.

Since the abstraction rules can access the alternation records, and since the alternation records are constrained by a language-independent set of roles shared by all translations of the verb being abstracted, the abstraction rules can identify the complements corresponding to each language-independent role and label them correctly where the original role assigned depended upon the source language.

Conveniently, during abstracting, the references from each term in the parsed structure to its source language word entry are replaced by references to the corresponding record in the language-independent meaning store 1243. At the same time, "register" or "tone" data indicating the tone (for example, "normal", "formal" or "informal") of each word entry is stored with the reference to the meaning entry. Several different word entries can correspond to the same meaning entry with different tones.

#### De-abstracting

In de-abstracting, the meaning entry references of the terms of the interlingual structure are each looked up in the meaning store 1243, and a word entry (in the target language word store 1242) with corresponding tone data to that stored for each term is selected. The roles present for each event term are then compared with those for each alternation record of the alternation class pointed to by the selected word entry, and the best- matching alternation record is selected.

The language-independent roles of the interlingual structure are then replaced, where necessary, as specified by that alternation (for example, where the verb in the target language has its roles reversed relative to the source language).

References to the selected target language word records are then substituted for the references to meaning records in the interlingual representation of the document.

#### Generation

Generation of the target language text then proceeds using the selected word records.

#### Generation of alternation data

The process of creating the data records used in the present embodiment will now be described.

Conveniently, the input and editing processes may be performed using the terminal 100 to access the server 200, from which the lexical database 240 and other records are read and to which they are written, via a browser program providing a graphical user interface into which data may be input and edited.

In a step 2002, the role set records (or, at any rate, most of them) are created by the user, and each meaning entry in the lexical database 240 which can have multiple alternations is assigned to one of the role sets (as mentioned above, there are typically 15-20 such role sets).

In a step 2004, a first language word store 1241 employed in the translation system (either as a source or a target language or both) is selected. For each language word store, the word entries will already have been assigned pointers to corresponding meaning entries in the meaning store 1243.

In a step 2006, a first event entry in the word store is selected.

Next, in a step 2008, the alternation classes associated with the role set assigned to that first event entry, in the language concerned, are displayed. If the event is the first event associated with that role set to be considered, there will be no alternation class displayed.

The data displayed (step 2016) for the alternation class is the list of alternations of the alternation class, displaying for each the role-complement mappings present in that alternation.

If no suitable class exists yet (step 2010), a new class is created (step 2012). Usually, a suitable class will exist already. In either case, in step 2014, the event is allocated to the alternation class it matches or the class which has newly been created.

If no alternations have yet been defined for the class, a template alternation listing the roles present in the class, in some order, is displayed and the user edits the display to re-order the roles into the desired order, add prepositions as desired, and so on.

If (step 2018) the list of alternations does not match those known by the inputter to exist for the word in the language concerned, then new alternations are created (step 2020) in the same way and added to the alternation class (step 2022).

If the last event in the language has not been reached (step 2024) the next event is selected (step 2026) and steps 2006 onwards are repeated.

If there are more languages to process (step 2028) the next is selected (step 2030) and the process returns to step 2004.

When all languages have been processed (or at any other desired end point) the data input is stored (step 2032) as alternation class records (702-706), alternation records 702-722, and role set records 730.

As mentioned above, it is typically found that relatively small number of role set records and a larger, but still small, number of alternation classes (of the order of a few hundred) per language required. The small number of role set records results from the relatively small number of different roles which can be played in events, and the relatively small number of alternation classes results from the same fact, and also from the tendency of many verbs to behave similarly.

The number of alternation records will vary from class to class and from language to language. The number of records will increase with the mutability of the word order in each language and with the irregularity of word orders between different verbs.

#### Role preference data

As stated above, associated with each of the role fields in the role set records 730 may be a role preference field.

For example, the lexical database 2040 may be hierarchically arranged, as described in PCT/GB98/03774 filed 16/12/98 priority 17/12/97, so that, for example an entry for "computer" points to a hierarchically higher entry for "electrical equipment" which in turn points to a hierarchically higher entry for "man made artefact" which in turn points to a hierarchically higher entry for "artefact" and thence to an entry for "entity".

Where the lexical database is hierarchically ordered in this manner (or, with greater difficulty, even where it is not organised in this manner) the preference field associated with each role may be set to point to a corresponding entry in the lexical database.

Thus, for example, certain types of activity are performed only by living creatures, and some only by people, so that the preference data for the "agent" role in these cases will be set respectively to point to the entries in the lexical database for "living creature" and "person".

Indirectly, through the hierarchical arrangement of the lexical database 240, the preference data therefore also points to all the hierarchically lower instances of those general classes which are stored in the lexical database.

The usefulness of such preference data is seen where the output produced by the parser is either incomplete or ambiguous. For example, if a given part of a document can be parsed to give two meanings, allocating different words or phrases to different roles, the ambiguity may be resolved during abstracting.

Each possible such parsed structure is matched to locate the corresponding alternation record, from which the presumed roles of each part of the parsed structure are determined. The role set record for the alternations is then examined, and it is determined whether the entities allocated to each role correspond to those specified in the preferences. The meaning for which the entities correspond more closely to the specified preferences is then selected as likelier to be correct.

Similarly, where incomplete input text is located by the parser so that complete parse cannot be performed, but nonetheless it is possible to locate for example a verb and preposition so that the correct role set record can be located, a comparison of the preference data stored for the roles in the role set record with the entries in the lexical database for the text surrounding the verb may suffice to complete the parse by allocating roles to the text present.

#### Other Embodiments and Variants

Other words than verbs can benefit from the invention; it may, for example, be used to compile multiple word entries for words which can change their form - e.g. adjectives which have an adverbial form. Each alternation record within a class can have a different syntactic category (e.g. adverb and adjective) and the record can thus be used to specify whether the derivation of a different word form can take place, and what the feature changes should be.

Although it is preferred to retain the role set records where they store role restriction data, if such data is not used in translation then the user-created role set records need not be present in the translator, being only used to derive the alternations consistently between different languages as described above, and, in this case, the alternations of an alternation class constitute a set of language-independent rules, and thus the set of language-dependent alternation classes for all the languages of the translator together constitute a means for storing sets of language-independent roles in accordance with the present invention.

Although the above embodiments accept a text document, a speech recognition front-end is also possible, or an image scanner with optical character recognition could be employed.

Although the above described embodiments describe a translation system, in which the target language text is generated, it will be understood that it would be possible with advantage to utilise the interlingual language structure generated for other purposes; for example, to provide a natural language front end or input routine for control of a computer or other equipment. Accordingly, such other uses of some aspects of the invention are not excluded.

Although adaptation to the intended target languages by limiting the search within the lexical database 240 to those words occurring in the source and those target languages has been described, it will be realised that it would also be possible to limit the operation of the abstractor, and merely to utilise those abstraction rules which remove language dependency in the source language which is not also present in the intended target languages.

In this case, each abstraction rule would similarly include a reference to those languages for which it was necessary, and only the necessary rules for the intended target language(s) would be used. Such an embodiment may prove useful as the number of target languages increases.

The foregoing embodiments are merely examples of the invention and are not intended to be limiting, it being understood that many other alternatives and variants are possible within the scope of the invention. Protection is sought for any and all novel subject matter disclosed herein and combinations of such subject matter.

Unless the context clearly requires otherwise, throughout the description and the claims, the words "comprise", "comprising" and the like are to be construed in an

•

## CLAIMS

1. Apparatus for translating text from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the apparatus comprising:

means for storing sets of language-independent roles;

means for storing, for each said language,

a respective set of language-dependent alternation classes, each alternation class being referenced to a said set of language-independent roles and comprising a set of alternations, and

a plurality of words, each word representing a said event and being referenced to an alternation class of its respective set of language-specific alternation classes,

each alternation of an alternation class comprising, in a respective order, the roles of said set of language-independent roles to which that class is referenced and listing the correspondence between those roles and complements by which each event referenced to that alternation class can validly be represented in that language;

means for locating, in said text, a phrase comprising one or more words representing an event, and the complements representing roles associated with that event;

means for representing said phrase in a language-dependent semantic structure;

and

means for replacing said language-dependent semantic structure with an indication of the language-independent roles represented by said complements using said alternations.

2. Apparatus according to claim 1, wherein said events are defined by a verb in said source language, and said means for replacing are arranged to locate a preposition occurring with said verb, and to replace said preposition with an appropriate role.



3. Apparatus according to either preceding claim, wherein multiple said events are referenced to a common said alternation class.

4. Apparatus according to any preceding claim, wherein multiple said alternation classes are referenced to a common said set of language-independent roles.

5. Apparatus according to any preceding claim, wherein the alternations of an alternation class constitute a set of language-independent rules, and the respective sets of language-dependent alternation classes for all the languages of the apparatus together constitute said means for storing sets of language-independent roles.

6. A method of setting up a machine translation system comprising:  
 a first stage of establishing, for each language, a respective plurality of language-dependent alternation classes, each alternation class comprising a plurality of alternations derived using a common set of language-independent role data shared by words of each language with a common meaning, and a respective set of word entries, each word entry being referenced to a said respective language-dependent alternation class; and

a second stage of creating, for each said alternation class and for each word entry referenced to that alternation class, a corresponding set of word entries identical to the word entry referenced to that alternation class, each word entry of said set being referenced to a different one of the alternations of that alternation class.

7. A method of translating text from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the method comprising the steps of:

storing sets of language-independent roles;

storing, for each said language,

a respective set of language-dependent alternation classes, each alternation class being referenced to a said set of language-independent roles and comprising a set of alternations, and

a plurality of words, each word representing a said event and being referenced to an alternation class of its respective set of language-specific alternation classes,

each alternation of an alternation class comprising, in a respective order, the roles of the set of language-independent roles to which that class is referenced and listing the correspondence between those roles and complements by which each event referenced to that alternation class can validly be represented in that language;

locating, in said text, a phrase comprising one or more words representing an event, and the complements representing roles associated with that event;

representing said phrase in a language-dependent semantic structure; and

replacing said language-dependent semantic structure with an indication of the language-independent roles represented by said complements using said alternations.

8. A method according to claim 7, wherein said events are defined by a verb in said source language, and said replacing step is arranged to locate a preposition occurring with said verb, and to replace said preposition with an appropriate role.

9. A method according to any either claim 7 or claim 8, wherein, in said step of storing a plurality of words, multiple said words representing a said event are referenced to a common said alternation class.

10. A method according to claim 9, wherein, in said step of storing a respective set of language-dependent alternation classes, multiple said alternation classes are referenced to a common said set of language-independent roles.

11. Apparatus for translating a document from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the apparatus being substantially as herein described with reference to the drawings.

12. A method of translating a document from a source language to an interlingual representation which can then be transformed into one or more of a plurality of target

languages, each of said languages including words corresponding to events and employing a concept of entities which play predetermined roles, the method being substantially as herein described with reference to the drawings.

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2224  
2225  
2226  
2227  
2228

# ABSTRACT

A method of abstracting a language-dependent source language semantic structure to provide a language-independent semantic structure, using abstracting rules defining a set of alternations which words of the source language can take, a plurality of said words sharing each said set.

1. A method of abstracting a language-dependent source language semantic structure to provide a language-independent semantic structure, using abstracting rules defining a set of alternations which words of the source language can take, a plurality of said words sharing each said set.

1/13

Fig.1.

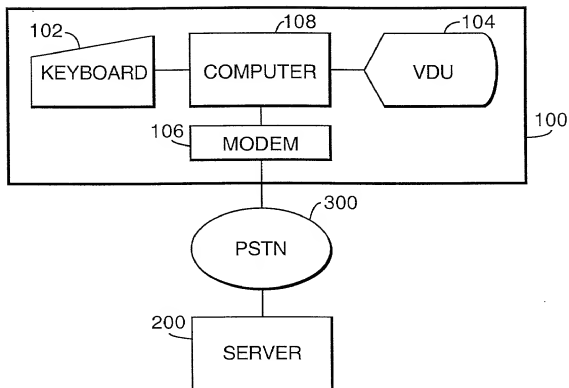
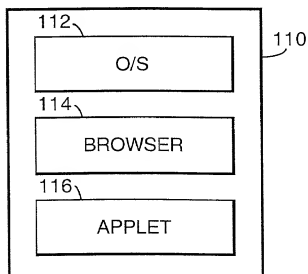


Fig.2.



2/13

Fig.3.

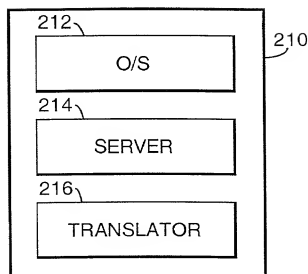


Fig.4.

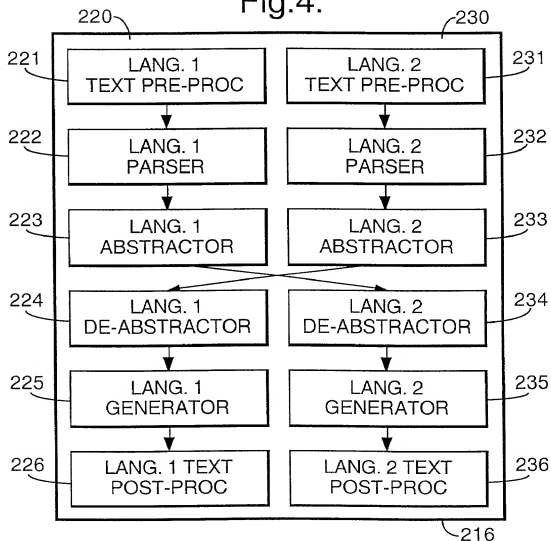


Fig.5.

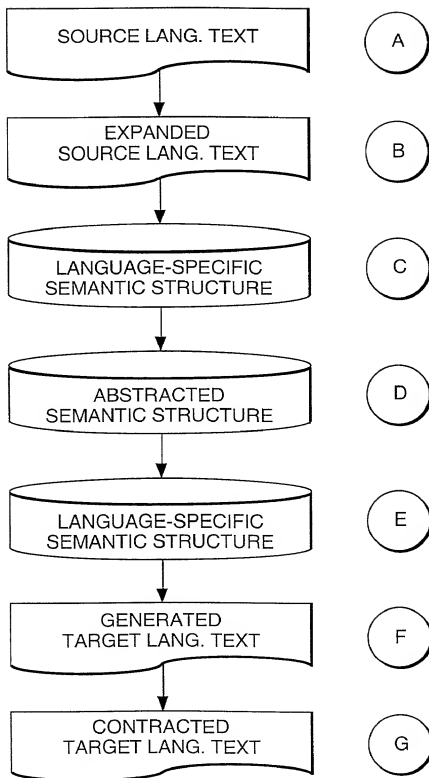


Fig.6.

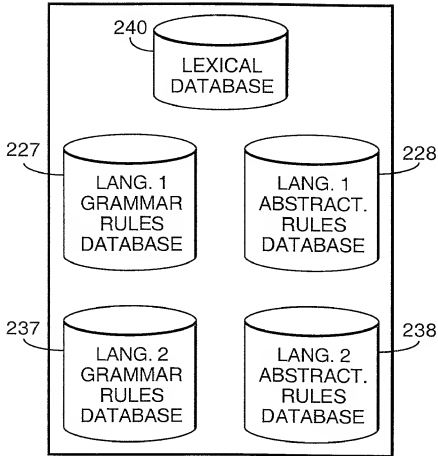
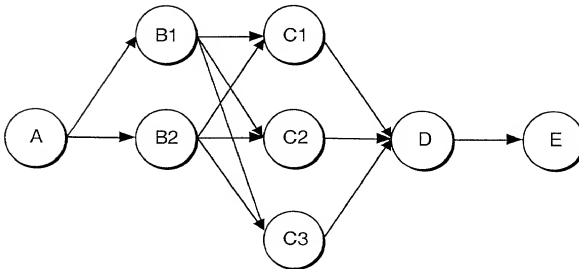


Fig.7.





5/13

Fig.8.

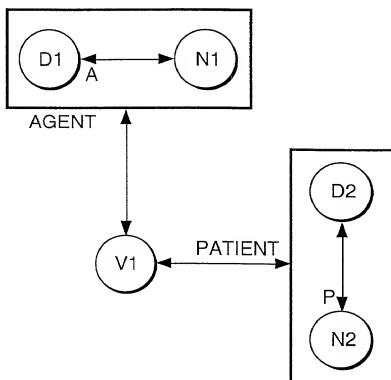
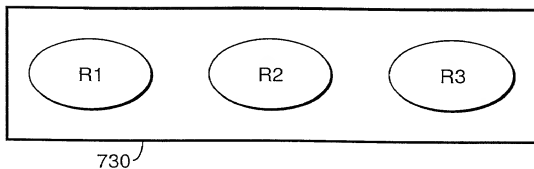


Fig.17.



6/13

Fig.9.

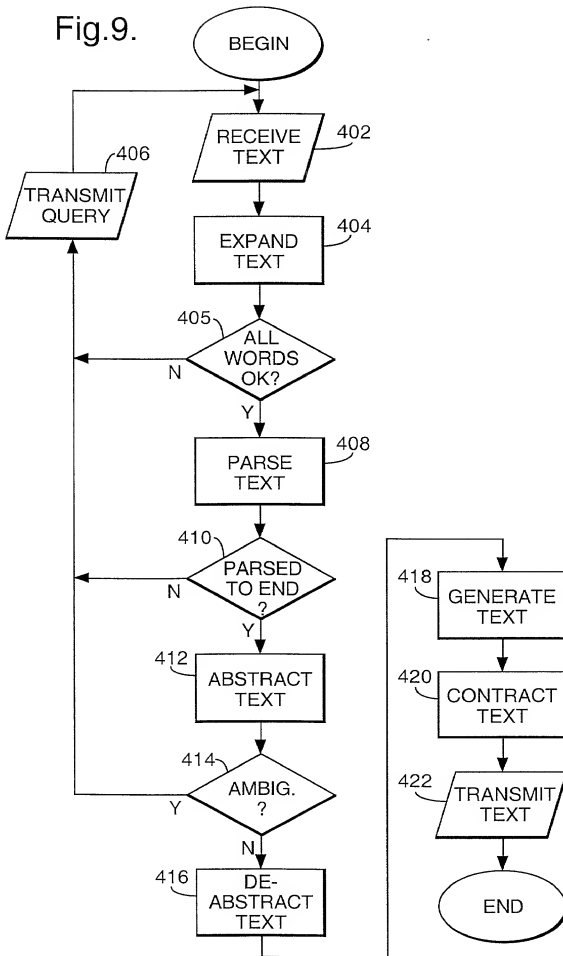


Fig.10.

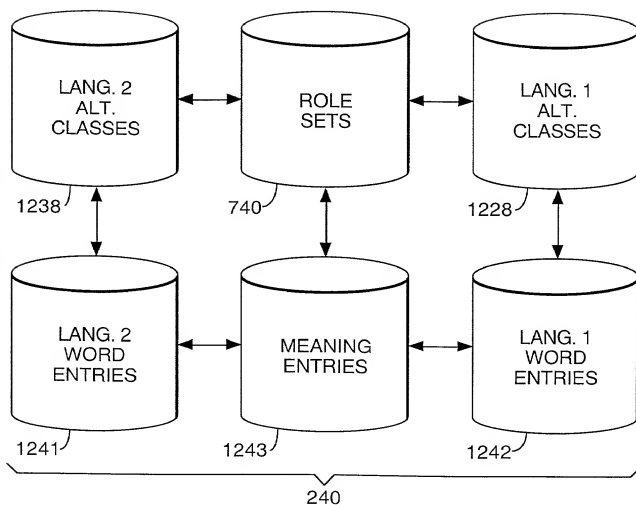
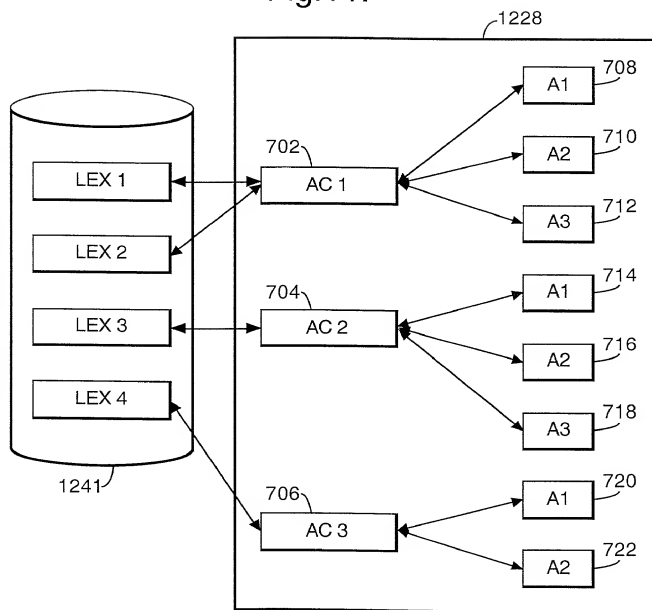


Fig.11.



9/13

Fig.12.

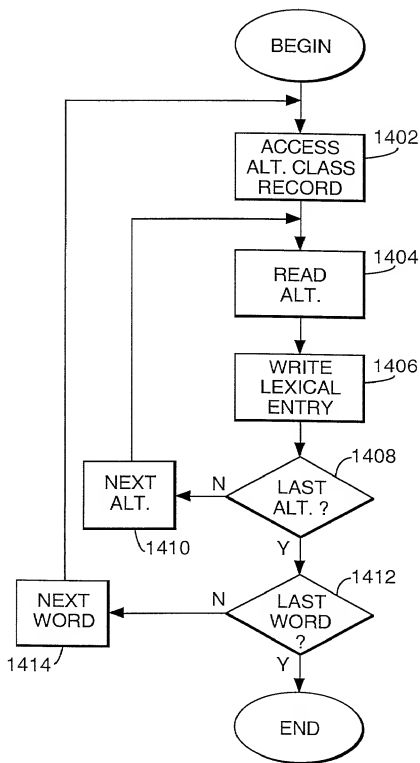


Fig.13.

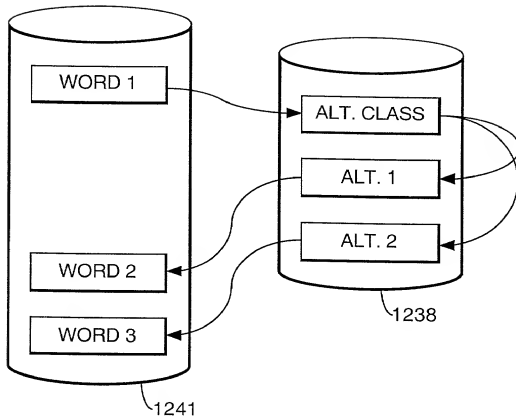


Fig.14a.

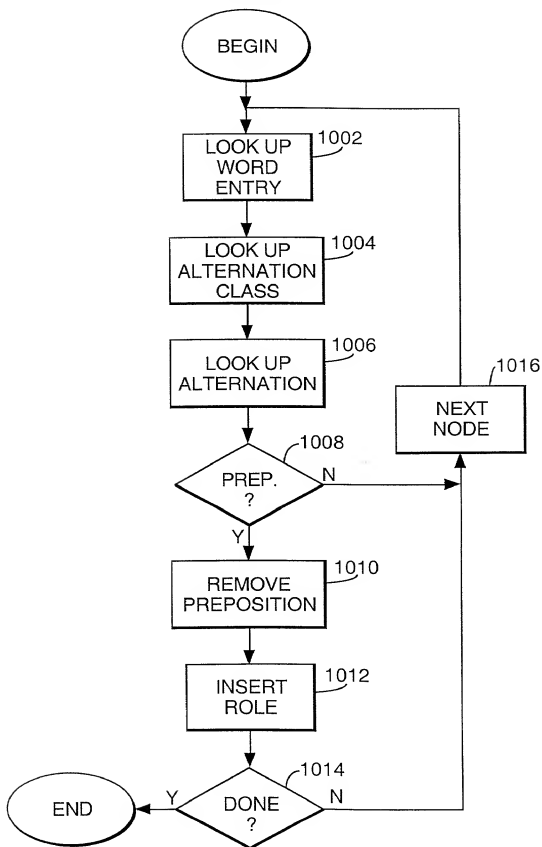


Fig.14b.



11/13

Fig.15.



12/13

Fig.16a.

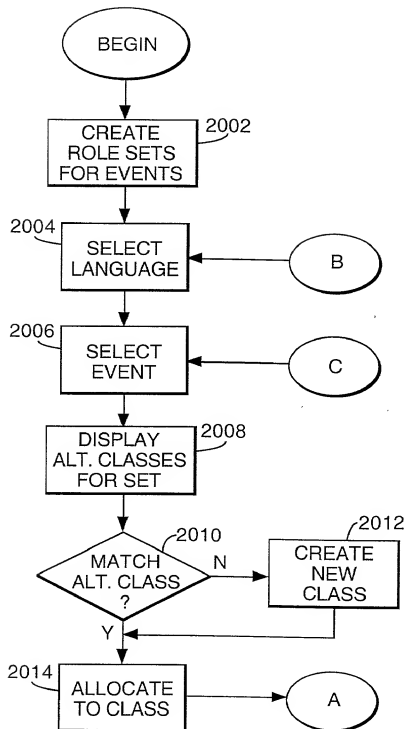
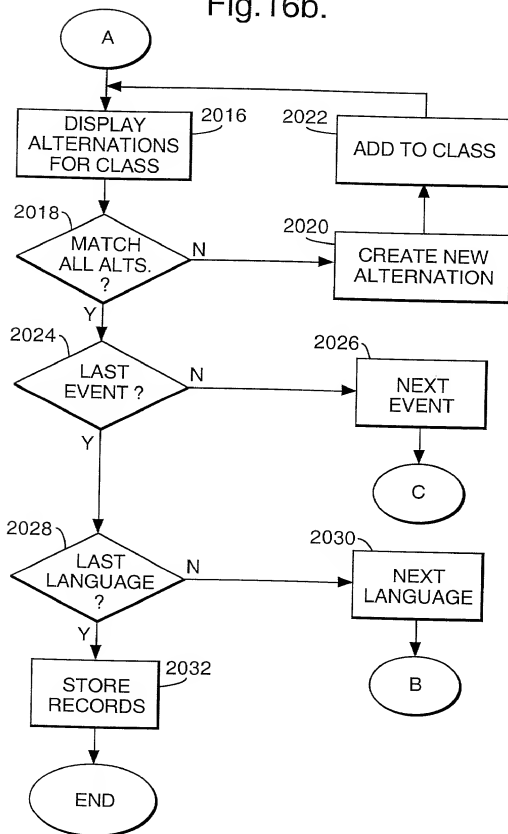




Fig.16b.



RULE 63 (37 C.F.R. 1.63)  
DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

As a below named inventor, I hereby declare that my residence, post office address and citizenship are as stated below next to my name, and I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

TRANSLATION

the specification of which (check applicable box(es)):

- ☐ is attached hereto  
☐ was filed on

as U.S. Application Serial No.

(Atty Dkt. No.)

☒ was filed as PCT International Application No.

PCT/GB00/00440 on 11 FEBRUARY 2000

and (if applicable to U.S. or PCT application) was amended on

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above. I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with 37 C.F.R. 1.56. I hereby claim foreign priority benefits under 35 U.S.C. 119/365 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed or, if no priority is claimed, before the filing date of this application:

Priority Foreign Application(s):

Application Number

9903747.5

Country

GREAT BRITAIN

Day/Month/Year Filed

18 FEBRUARY 1999

9904784.4

EUROPE

18 JUNE 1999

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

Application Number

Date/Month/Year Filed

I hereby claim the benefit under 35 U.S.C. 120/365 of all prior United States and PCT international applications listed above or below and, insofar as the subject matter of each of the claims of this application is not disclosed in such prior applications in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose material information as defined in 37 C.F.R. 1.56 which occurred between the filing date of the prior applications and the national or PCT international filing date of this application:

Prior U.S./PCT Application(s):

Application Serial No.

Day/Month/Year Filed

Status: patented

pending, abandoned

PENDING

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon. And on behalf of the owner(s) hereof, I hereby appoint NIXON & VANDERHYE P.C., 1100 North Glebe Rd., 8th Floor, Arlington, VA 22204-4714; telephone number (703) 815-4000 (to whom all communications are to be directed), and the following attorneys thereof (of the same address) individually and collectively owner's/owners' attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith and with the resulting patent: Arthur R. Crawford, 25327, Larry S. Nixon, 25640; Robert A. Vanderhye, 27076; James T. Hosmer, 30184; Robert W. Farris, 31352; Richard G. Besha, 22770; Mark E. Nusbaum, 32348; Michael J. Keenan, 32106; Bryan H. Davidson, 30251; Stanley C. Spooner, 27393; Leonard C. Mitchard, 29008; Duane M. Byers, 33363; Jeffrey H. Nelson, 30481; John R. Lastova, 33145; H. Warren Burnham, Jr., 29366; Thomas E. Byrne, 32205; Mary J. Wilson, 32955; J. Scott Davidson, 33489; Alan M. Kagen, 36178; Robert A. Molan, 29834; B. J. Sadoff, 33663; James D. Berquist, 34726; Updeep S. Gill, 37334; Michael J. Shea, 34725; Donald L. Jackson, 41090; Michelle N. Lester, 32331; Frank P. Presta, 18828; Joseph S. Presta, 38322. I also authorize Nixon & Vanderhye to delete any attorney names/numbers no longer with the firm and to act and rely solely on instructions directly communicated from the person, assignee, attorney, firm, or other organization sending instructions to Nixon & Vanderhye on behalf of the owner(s).

Inventor's Signature:

Inventor:

STEPHEN

(first)

MI

Date:

6

March

2000

GB

Residence: (city)

ESSEX

GBN

(state/country)

GREAT BRITAIN

Post Office Address:

23 RICHARD AVENUE, WIVENHOE, COLCHESTER, ESSEX

(Zip Code)

CO7 9JQ

2.

Inventor's Signature:

Inventor:

(first)

MI

(last)

(citizenship)

Residence: (city)

(state/country)

Post Office Address:

(Zip Code)